

LA RECONNAISSANCE DU PATTERN PROSODIQUE DE LA QUESTION :

questions de méthode.

Monique Vion & Annie Colas

Université de Provence
CNRS UMR 6057 : Parole et Langage, équipe Psycholinguistique
mvion@newsup.univ-mrs.fr

Résumé

L'étude linguistique de la forme prosodique de la question en français montre que dans les questions comportant plus de deux unités tonales, les unités tonales intermédiaires présentent un pattern caractéristique d'abaissements successifs de la hauteur.

Le paradigme du dévoilement graduel est utilisé pour montrer que les auditeurs sont capables d'identifier le caractère interrogatif d'un énoncé au cours du dévoilement de l'unité tonale intermédiaire de phrases nominales composées de trois unités.

Les résultats montrent que, bien que la tâche soit extrêmement difficile, il est possible aux auditeurs de reconnaître une question sur la base des caractéristiques prosodiques de l'unité tonale intermédiaire, mais qu'ils n'en sont que moyennement sûrs.

Abstract

Linguistics studies of intonational devices of total questions in French showed that, in questions containing more than two stress groups, interrogative intonation is characterized by a sequence of lowered pitches or downstepped tones.

The gating paradigm was used to determine whether subjects who are listening to the mid stress group of a French NP utterance containing three stress groups can indicate whether the utterance is an affirmation or a question.

Although the task was extremely difficult, findings indicate that listeners are infrequently but effectively able to recognize, with mid confidence ratings, the intonational device of a question while they are hearing the end part of the mid stress group.

Mots clés : Français, intonation de la question, paradigme du dévoilement graduel.

Key words : French language, interrogative intonation, gating paradigm.

L'article présente une recherche en cours où s'articulent les travaux issus de plusieurs disciplines, au service de l'approche d'une dimension du langage longtemps négligée tant en linguistique qu'en psycholinguistique, mais pourtant essentielle : la prosodie. L'étude porte plus précisément sur le rôle des indices prosodiques dans le traitement de la modalité interrogative en français.

Di Cristo & Hirst (1993) se sont attachés à décrire les régularités prosodiques caractéristiques de différents types de questions en français. Cette étude de la structure de surface des questions s'inscrit dans le cadre d'une approche linguistique qui vise à définir les structures prosodiques au plan phonologique (pour une présentation synthétique de la théorie de l'intonation développée dans : Hirst et Di Cristo, 1984 ; Hirst, 1988 ; Hirst et Di Cristo, 1998 ; voir Di Cristo, 2000).

Quatre types de questions composées d'une ou de deux unités intonatives ont été étudiés : les questions qui appellent une réponse par « oui » ou par « non » (« tu veux du fromage ? ») ; les questions elliptiques (« et du dessert ? ») ; celles qui proposent un choix alternatif (« tu veux du fromage ou du dessert ? ») et celles qui comportent un détachement à gauche (« du dessert, tu en veux aussi ? »). L'analyse de corpus de parole spontanée (enregistrements d'entretiens radiophoniques) ainsi que celle de courts dialogues simulés a conduit les auteurs à contraster deux types de contours finaux pour ces questions. Alors que les questions qui appellent une réponse par oui ou par non ainsi que le premier terme des questions alternatives et la partie rhématique des questions qui comportent un détachement à gauche présentent un contour final ascendant, les questions elliptiques ainsi que le deuxième terme des questions alternatives et la partie thématique des questions à détachement à gauche, présentent un contour final descendant. En outre, lorsque les unités intonatives comportent plus de deux unités tonales et quel que soit le type de la question, la description d'ensemble du pattern de hauteur, montre la présence d'abaissements successifs de la hauteur pour les unités tonales intermédiaires.

Ce pattern d'abaissements successifs de la hauteur pour les unités tonales intermédiaires, pour ne pas être seulement lié à un contour ascendant subséquent, constituerait selon les auteurs une propriété phonologique de la question en français et pas seulement une caractéristique liée à l'anticipation du mouvement subséquent.

Le but de la présente recherche est d'apprécier la valeur distinctive de ces abaissements successifs de la valeur de la fréquence fondamentale sur les unités tonales intermédiaires. Un auditeur peut-il identifier l'énoncé d'une

phrase comme une question avant d'avoir entendu la dernière unité tonale, porteuse du contour final ? Autrement dit, les abaissements successifs de la hauteur sur les unités tonales intermédiaires d'une phrase énoncée constituent-ils des indices pour l'auditeur qui lui permettent d'en identifier (en l'anticipant) la modalité ?

Dans la palette des paradigmes expérimentaux exploités en psycholinguistique, le choix du dévoilement graduel (*gating*) semblait approprié pour répondre à cette question. Ce paradigme est en effet pertinent pour tout un ensemble de questions dont celles de la détermination de la quantité d'information acoustico-phonétique nécessaire pour identifier un stimulus ou du rôle joué par les variables, phonétiques, lexicales et contextuelles au cours de l'identification (Grosjean, 1996). Il est à noter que ce paradigme a été peu utilisé jusqu'ici pour l'étude du rôle des indices prosodiques dans le traitement des énoncés (Grosjean, 1980b ; Grosjean et Hirt, 1996), son domaine d'application le plus fréquent étant celui de l'identification lexicale. C'est la raison pour laquelle les difficultés rencontrées par les participants lors de son emploi dans la présente recherche méritaient d'être exposées.

Dans la version remise à l'honneur par Grosjean (1980 a et b), une série de stimulus comportant des portions de plus en plus longues d'un élément de langage parlé (syllabe, mot, syntagme ou phrase) est constituée en partant à chaque fois du début du signal. Le premier stimulus est très bref et le dernier correspond à la totalité de l'élément considéré. Le pas d'incrémentation peut être exprimé en millisecondes ou en pourcentage de l'élément dévoilé. L'auditeur à qui les stimulus sont proposés a pour tâche de les identifier successivement et d'exprimer sa confiance en chaque réponse donnée. Les réponses recueillies permettent : 1) de déterminer un point d'isolement (à savoir, la taille du signal dévoilé nécessaire à l'identification de l'élément étudié) ; 2) de connaître le degré de certitude de l'auditeur en tout point du dévoilement (notamment au point d'isolement et au dernier stimulus) ; et 3) d'analyser les caractéristiques des candidats proposés antérieurement à l'identification.

Les modalités de phrase sont liées à des marques formelles caractéristiques : le mode verbal, l'ordre sujet verbe et l'intonation de fin de phrase (« où es-tu ? »). Un deuxième choix semblait pertinent pour tenter de répondre à la question posée : étudier un type de phrase où l'intonation joue un rôle décisif dans le marquage formel de la modalité : les phrases nominales.

Rencontrées parfois sous forme de formules figées (« au diable l'avarice ! », « gloire à dieu au plus haut des cieux et paix sur la terre aux hommes de bonne volonté. »), les phrases nominales ou non-verbales sont très courantes dans l'expression orale spontanée (« très bien, ton exposé. »). On

les considère généralement comme le fruit d'une réaction immédiate des interlocuteurs, commandée par la dynamique de la communication.

D'un point de vue grammatical, bien qu'étant des phrases sans verbe ce sont des phrases complètes dans la mesure où elles se composent de deux termes, un sujet (ton exposé) et un prédicat (très bien), mis en relation par l'énonciateur (Le Goffic, 1993). Les moyens syntaxiques mis en œuvre pour prédiquer au moyen de ces phrases, sont la prosodie et l'ordre des mots. Comme on peut le voir dans les exemples ci-dessus, l'absence de verbe n'empêche pas la manifestation d'une modalité de phrase (assertions ou exclamations le plus souvent).

D'un point de vue pragmatique, les phrases nominales interrogatives semblent plutôt destinées à appeler une réponse par «oui» ou par «non». Il peut s'agir d'une demande de confirmation (A: «je me suis inscrit en troisième année de droit», B: «en troisième année de droit ?») ou d'une reprise dubitative d'une assertion (A: «tu as réussi ton examen», B: «j'ai réussi mon examen ?»).

Le type de phrases nominales retenu pour l'étude est celui dans lequel un terme nominal représente un prédicat d'existence (phrases nominales existentielles). Le groupe nominal, constituant unique, incorpore ce qui pourrait être développé en une phrase. Ce troisième choix avait l'avantage d'envisager des phrases pouvant aller d'un mot unique («imbécile !») à une suite de mots dont la longueur est seulement limitée par la capacité de la mémoire immédiate (A: «tu fais quoi, toi, dans la vie ?», B: «adjoint du directeur du centre de recherche géologique de l'institut régional de diffusion de la science de l'académie de Montpellier») et donc de construire des stimulus pouvant comporter un nombre variable d'unités tonales.

Pour amorcer l'étude expérimentale, un sondage préliminaire a été effectué où les auditeurs avaient à traiter des phrases nominales constituées de deux («les paquets du livreur», absence d'unité tonale intermédiaire), de trois («la couleur des cyprès du jardin», une unité tonale intermédiaire) et de quatre unités tonales («le tissu des coussins du fauteuil du salon», deux unités tonales intermédiaires). Le pas de dévoilement choisi était l'unité tonale. Les auditeurs devaient porter un jugement en choix forcé sur chaque stimulus proposé. Ils devaient dire si le stimulus entendu provenait de l'énoncé d'une phrase affirmative ou d'une phrase interrogative. L'analyse des résultats de ce sondage a montré la difficulté des auditeurs, hors du contexte conversationnel, à traiter un énoncé sans verbe comme une phrase achevée. En effet, la modalité interrogative n'a pas toujours été identifiée à 100% comme telle lorsqu'un stimulus était entièrement dévoilé. De même, bien que rarement, il est arrivé qu'un énoncé affirmatif entièrement dévoilé ne soit pas reconnu. Quand un énoncé interrogatif a été reconnu comme tel, la certitude exprimée alors que le stimulus était entièrement dévoilé était de l'ordre de

90%. Par ailleurs 10% des auditeurs environ ont identifié la modalité interrogative du stimulus à l'issue du dévoilement des unités tonales intermédiaires. Cette difficulté de traitement des phrases nominales était renforcé en partie par le mode de réponse exigé qui demandait de répondre de mémoire, sans que les phrases alternatives écrites et ponctuées ne soient proposées en support. La présence de réponses pertinentes antérieures à l'audition complète des énoncés incitait cependant à explorer le déroulement temporel de ces identifications précoces.

L'expérience présentée ci-dessous visait à étudier l'évolution des réponses données au cours de l'audition de l'unité tonale intermédiaire de phrases nominales comportant trois unités tonales. Le mode de réponse était aménagé de telle sorte que les participants à l'expérience disposent, à l'issue de l'audition du stimulus, du support écrit des phrases alternatives pour effectuer leurs jugements.

METHODE

Matériel

Phrases

Les phrases qui ont servi à constituer les stimulus test et les stimulus d'entraînement sont des phrases nominales, reprises de Soulier (1995). Toutes les phrases sont constituées de trois unités tonales. Chaque unité tonale comporte trois syllabes de type Consonne voisée / Voyelle, de façon à conserver une ligne mélodique ininterrompue (voir la liste des phrases en annexe 1).

Chaque phrase a été enregistrée deux fois par un locuteur féminin : une fois avec l'intonation caractéristique de l'affirmation et une fois avec l'intonation caractéristique de la question. Les enregistrements ont été soumis à un expert phonéticien afin de s'assurer d'une part, que la façon de prononcer les phrases affirmatives et interrogatives était homogène et d'autre part, que l'intonation des phrases était conforme aux patterns décrits dans la littérature à savoir pour l'affirmation : "*a basic pitch pattern (ie an Intonation Unit) ending on a final low pitch (such as a statement) contains a rising pitch movement (from low to high) at the end of each stress-group, except the last which is pronounced with a falling pitch movement from the mid (or lowered high) to low.*" (Di Cristo (1998), p. 201; et pour la question : "*an initial peak situated on the last syllable of the first stress-group followed by a declining pitch until the final rise.*" (Di Cristo (1998), p. 203 (Tableau 1).

Tableau 1 : Variations de hauteur au sein des unités tonales (d'après Di Cristo, 1998).

L'aMI du voiSIN de JEAN

Assertion [↑ ↓ ↑ > ↓↓]

L'aMI du voiSIN de JEAN

Question [↑ > > ↑↑]

Légende : La partie accentuée est représentée en capitales ; ↓ : "plus bas" ; ↑ : "plus haut" ; > : abaissement ; ↑↑ : "haut" (hauteur maximale) ; ↓↓ : "bas".

La figure 1, qui présente l'analyse acoustique de l'une des phrases du matériel, prononcée sous les deux formes, montre les variations de la fréquence fondamentale caractéristique de chaque pattern. La valeur de la fréquence fondamentale mesurée sur la voyelle de la sixième syllabe des phrases du matériel est en moyenne de 240,95 Hz pour les phrases affirmatives et de 191,49 Hz pour les phrases interrogatives.

Figure 1a

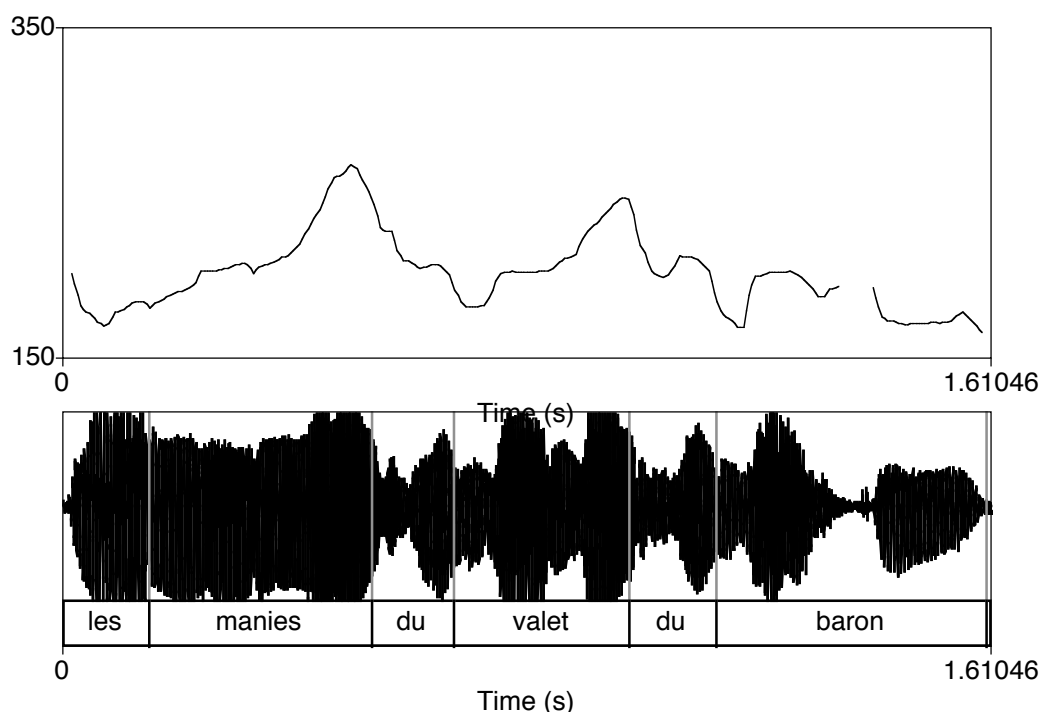


Figure 1b

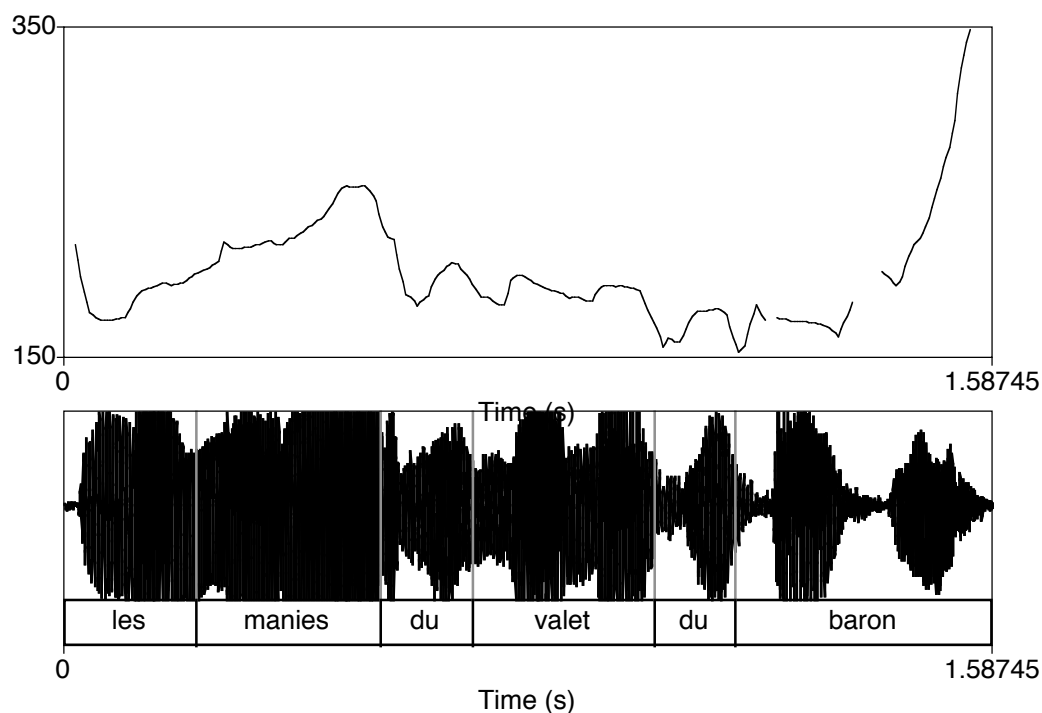


Figure 1 : Analyse acoustique de la phrase « les manies du valet du baron » prononcée en (a) comme une assertion et en (b) comme une question.

Haut : courbe des variations de la F0

Bas : signal acoustique

Stimulus

Pour constituer les stimulus, les trois syllabes qui portent le contour final ont été supprimées des enregistrements. Pour chaque enregistrement ainsi tronqué (où ne restent donc que les six premières syllabes) un ensemble de six stimulus (G : *gates*) a été constitué. Le premier stimulus (G1) comporte les 3 syllabes de la première unité tonale. Les stimulus suivants comportent, outre les 3 syllabes de G1 un fragment progressivement de plus en plus long des syllabes qui constituent la seconde unité tonale (le pas d'accroissement est de 20% : les stimulus G2, G3, G4, et G5 comportent respectivement 20%, 40%, 60% et 80% de la seconde unité tonale). Le dernier stimulus comporte la totalité des six syllabes (G6 : 100% de la seconde unité tonale) (figure 2).

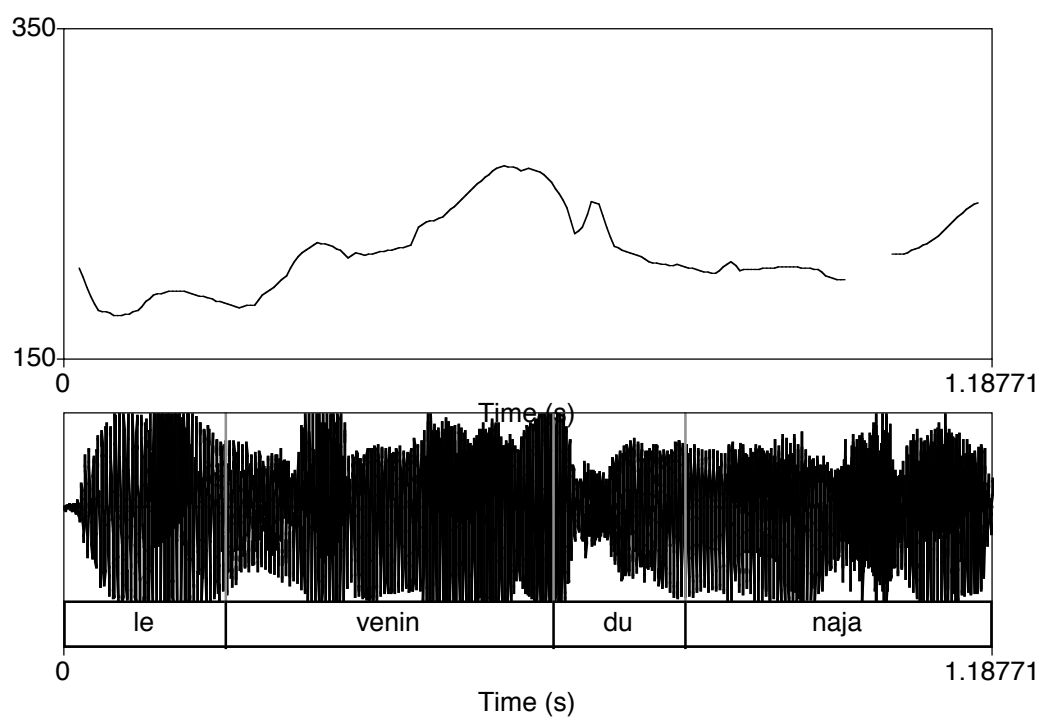


Figure 2b

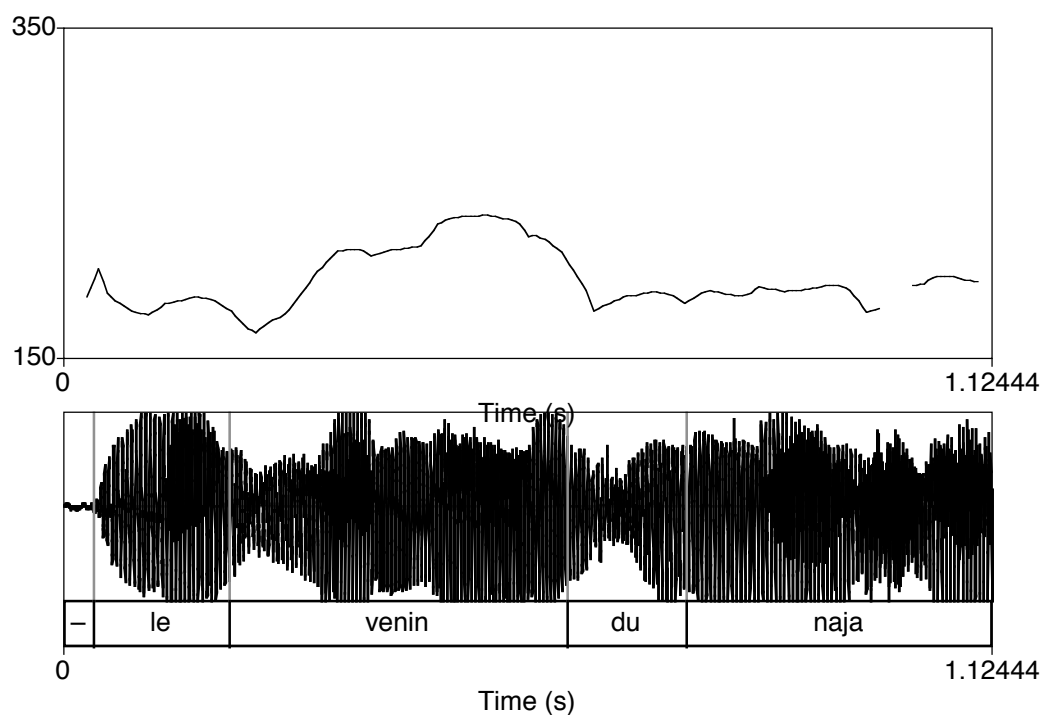


Figure 2 : Analyse acoustique des deux stimulus G6 issus de la phrase «le venin du naja d'Ouganda» prononcée en (a) comme une assertion et en (b) comme une question.

Haut : courbe des variations de la F0

Bas : signal acoustique

Items d'entraînement

Pour familiariser l'auditeur avec la tâche, une liste de trois items (soit 3 ensembles de 6 stimulus) a été constituée sur le modèle décrit ci-dessus de façon à ce qu'il entende une fois l'intonation de l'affirmation et une fois celle de la question. Un ensemble de stimulus issu d'une phrase énoncée avec l'intonation de l'implication est également présenté. L'ordre de présentation des trois items est constant pour tous les participants, soit : affirmation, implication, question.

Items test

Deux listes ont été constituées pour les seize items tests de façon à ce qu'un contenu donné soit entendu : avec l'intonation de l'affirmation dans une liste et avec celle de la question dans l'autre liste. Dans chaque liste, les seize contenus sont donc présentés ou bien sous forme d'affirmation (8 items) ou bien sous forme de question (8 items).

Procédure

La passation, pilotée par ordinateur, est individuelle. L'expérimentateur invite tout d'abord l'auditeur à lire à haute voix la consigne au fur et à mesure de son affichage sur l'écran. L'expérimentateur la délivre pas à pas en cliquant sur la souris et la commente si besoin est, page d'écran après page d'écran (voir la consigne en annexe 2). Puis il invite l'auditeur à placer les écouteurs qui lui permettront d'entendre les stimulus.

L'auditeur passe ensuite les trois items d'entraînement, puis les 16 items de l'expérience proprement dite. Le programme enregistre la réponse d'identification et la certitude qui l'accompagne, ainsi que le temps mis pour donner chacune des deux réponses. L'expérience dure en moyenne trente minutes.

Les essais d'entraînement ainsi que ceux de l'expérience proprement dite ont la même structure. Un stimulus est d'abord diffusé. Immédiatement après la fin de celui-ci, une première page d'écran invite à donner au clavier la réponse d'identification en choix forcé (« A » ou « Q »). Simultanément les deux versions possibles de la phrase s'affichent l'une au dessous de l'autre dans la partie haute de l'écran (l'affirmation est indiquée par un point final et la question par un point d'interrogation). Chaque phrase est en outre précédée de l'étiquette A ou Q respectivement. Une fois la réponse d'identification saisie, une seconde page d'écran s'affiche. Elle invite à exprimer (en appuyant sur l'une des 9 touches du pavé numérique) le degré de confiance en la réponse donnée (1 signifie : "pas sûr du tout" et 9 : "tout à fait sûr"). Une troisième page d'écran invite alors à passer au stimulus suivant en appuyant sur la barre d'espacement. En fin d'item, une page d'écran indique la fin de ce dernier et invite à passer à l'item suivant en appuyant sur la barre d'espacement.

Quatre programmes, écrits grâce au logiciel PsyScope (Cohen, Mac Whinney, Flatt et Provost, 1993), gèrent l'ensemble du recueil. Pour chacun, les affirmations et les questions sont présentées dans un ordre tiré au hasard sur l'ensemble des 16 items. Les programmes se différencient par la liste d'items et par l'ordre d'affichage des phrases à l'écran lors de l'invitation à répondre. Pour un programme donné, la moitié des participants dispose de la touche de réponse étiquetée « A » à gauche du clavier et de celle étiquetée « Q » à droite. Pour l'autre moitié la disposition des touches est inversée.

Facteurs

Les variables indépendantes manipulées sont : *l'intonation du stimulus* (facteur expérimental à deux modalités : affirmative (IA) et interrogative (IQ)) ; *la quantité d'information entrante* (facteur expérimental à six modalités : G1 : première unité tonale ; G2 : 20 % de la 2^{ème} unité tonale ; G3 : 40 % de la 2^{ème} unité tonale ; G4 : 60 % de la 2^{ème} unité tonale ; G5 : 80 % de la 2^{ème} unité tonale ; G6 : 100% de la 2^{ème} unité tonale).

Sont maintenues constantes les variables qui concernent : *l'identité du locuteur* (voix féminine) ; le *nombre d'unités tonales des phrases* (trois unités) ; *la taille des unités tonales* (3 syllabes).

A ces deux ensembles de variables, s'ajoutent des variables de contrôle : *la liste d'items*, qui contrôle les éventuels effets liés au contenu des phrases ; *l'ordre d'affichage des phrases à l'écran*, qui contrebalance les éventuels effets d'ordre d'apparition (au-dessus – au-dessous) des phrases (facteur à deux modalités) ; *l'ordre des touches de réponse* sur clavier qui contrebalance les éventuels effets d'ordre (Gauche - Droite) de disposition des touches de réponse (facteur à deux modalités).

Plan

Chaque participant est confronté aux *deux intonations*. Il entend chaque contenu sous une seule des intonations possibles. Ce qui implique la constitution de deux groupes de participants pour que chaque contenu ait été présenté sous ses deux intonations dans l'expérience. Chaque participant est confronté à un seul *ordre d'affichage des phrases à l'écran* qui l'invite au choix forcé. Ce qui implique la constitution de deux groupes. Chaque participant est confronté à une seule *disposition des touches de réponse au clavier*. Ce qui implique la constitution de deux groupes. La passation de l'expérience nécessite donc au total la constitution de huit groupes de participants passant chacun dans l'une des conditions expérimentales issues du croisement des caractéristiques : liste 2 x écran 2 x clavier 2. Quatre auditeurs passent dans chaque condition.

Participants

Trente deux jeunes adultes ayant le français comme langue première, étudiants de l'Université de Provence ont participé à l'expérience.

Attentes

Au plan phonologique la structure prosodique des phrases affirmatives et interrogatives ne se différencie qu'à partir de la seconde unité tonale. Sous l'hypothèse que pour les phrases comportant plus de deux unités tonales, l'auditeur peut commencer à différencier le pattern de la question de celui de l'affirmation dès l'audition de l'unité tonale intermédiaire, les attentes concernent l'évolution des réponses données au cours du dévoilement graduel de l'information prosodique supposée discriminante. Si au cours du dévoilement de l'unité tonale intermédiaire des phrases de l'expérience, les réponses données sont liées au traitement de l'information prosodique, la question devrait être de mieux en mieux reconnue au fur et à mesure qu'augmente la taille de la portion dévoilée de cette partie du signal.

RESULTATS

Chaque participant ayant donné 96 réponses d'identification (16 items x 6 essais) 3072 réponses d'identification (32 participants x 96) et 3072 réponses exprimant la certitude associée à chacune de ces réponses sont disponibles pour l'analyse. Les temps mis pour donner chaque réponse sont également disponibles.

Résultats globaux

Pour apprécier le niveau de difficulté de la tâche proposée, dont les stimulus ne présentent jamais la phrase achevée, les données ont d'abord été examinées du point de vue du temps mis pour donner une réponse d'identification, quelle que soit la nature de cette réponse. L'analyse de variance conduite avec les facteurs *2Intonation* x *6Quantité d'information entrante*, et le temps moyen de réponse aux stimulus pour variable dépendante, révèle au seuil choisi de .05, un effet significatif de la quantité d'information entrante ($F(5, 155) = 19,744$; $p < .00001$)(figure 3).

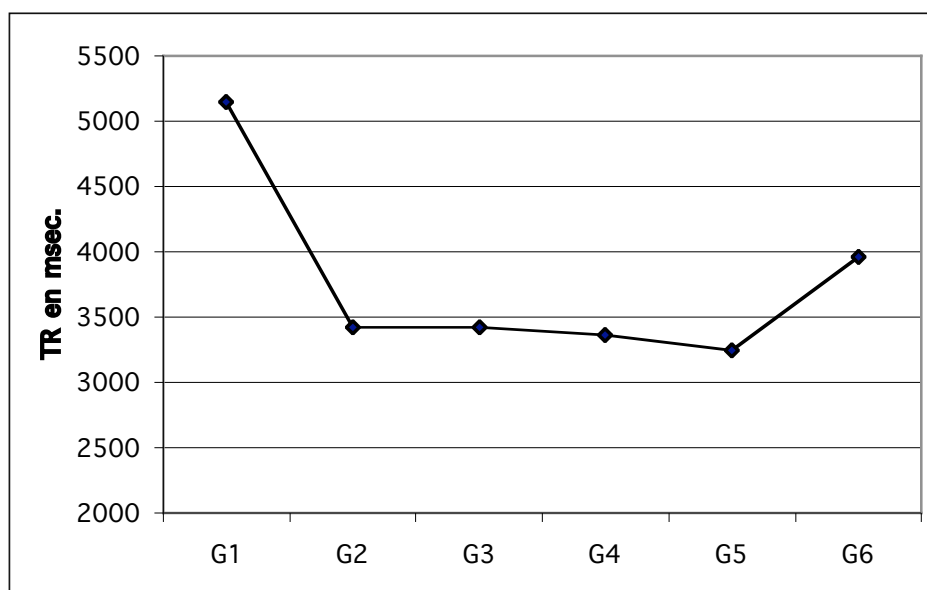


Figure 3 : Evolution du temps moyen mis pour donner une réponse d'identification au cours des essais.

Légende : G1 : première unité tonale ; G2 : 20 % de la 2^{ème} unité tonale ; G3 : 40 % de la 2^{ème} unité tonale ; G4 : 60 % de la 2^{ème} unité tonale ; G5 : 80 % de la 2^{ème} unité tonale ; G6 : 100% de la 2^{ème} unité tonale.

Quelle que soit l'intonation du stimulus, les temps de réponse qui sont de l'ordre de 5 secondes lors de la réponse à la présentation du premier stimulus d'une série (1^{ère} unité tonale), décroissent et se stabilisent aux environs de 3 secondes entre le second stimulus (20% de la 2^{ème} unité tonale) et le 5^{ème} stimulus (80 % de la 2^{ème} unité tonale) présentés. Les temps de réponse tendent à s'allonger à nouveau lors de la réponse à la présentation du dernier stimulus de la série (environ 4 secondes).

Deux effets conjugués peuvent rendre compte de ce résultat. D'une part un effet de familiarisation avec la série qui permet à l'auditeur de réduire son temps de réponse au fur et à mesure de la présentation des stimulus. Et d'autre part un effet lié au fait que certains stimulus (G1 et G6), qui donnent à entendre une (ou des) entité(s) complète(s) au plan segmental, incitent l'auditeur à accorder plus de soin à l'établissement de son jugement avant de donner sa réponse.

Les données ont été examinées ensuite du point de vue de la pertinence des réponses fournies. Pour cela, une variable dépendante a été dérivée des réponses d'identification « A » et « Q » saisies au clavier. Les réponses « A » données pour un stimulus affirmatif et les réponses « Q » données pour un stimulus interrogatif ont été considérées comme des réponses pertinentes (codées 1) ; les réponses « A » données pour un stimulus interrogatif et les réponses « Q » données pour un stimulus affirmatif ont été considérées comme des réponses non-pertinentes (codées 0). Une analyse de variance selon le plan décrit ci-dessus, conduite avec cette

variable dépendante (dont l'étendue varie de 0 à 8), montre au seuil choisi de .05, un effet significatif des facteurs *Intonation* ($F(1, 155) = 14,847$; $p < .0005$) et *Quantité d'information entrante* ($F(5, 155) = 7,468$; $p < .00001$) ainsi qu'un effet d'interaction entre ces deux facteurs ($F(5, 155) = 2,722$; $p < .02$) (figure 4).

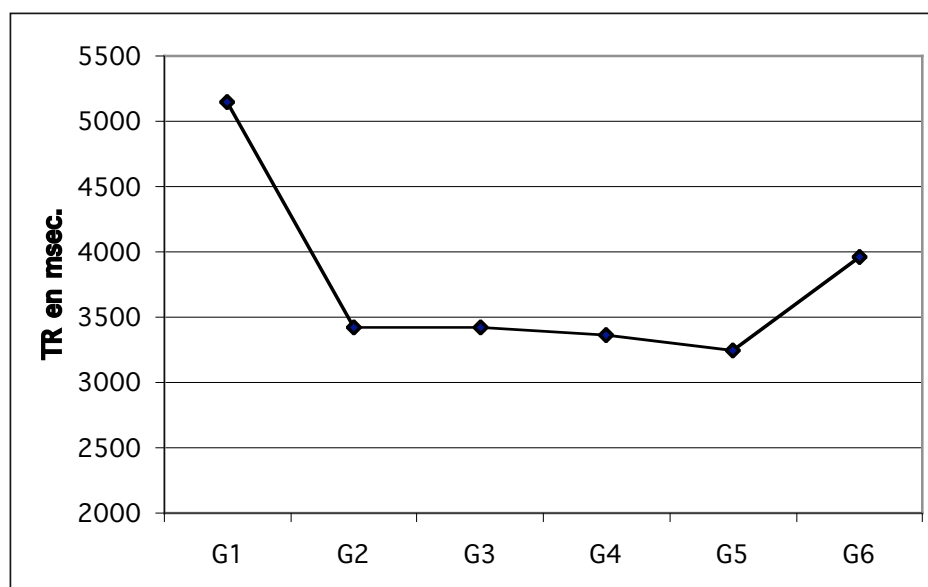


Figure 4 : Evolution du nombre moyen de réponses pertinentes au cours des essais.

Légende : G1 : première unité tonale ; G2 : 20 % de la 2^{ème} unité tonale ; G3 : 40 % de la 2^{ème} unité tonale ; G4 : 60 % de la 2^{ème} unité tonale ; G5 : 80 % de la 2^{ème} unité tonale ; G6 : 100% de la 2^{ème} unité tonale.

Quelle que soit l'intonation du stimulus, les réponses pertinentes augmentent régulièrement au fur et à mesure du dévoilement. Elles sont à leur minimum à l'issue de la présentation du premier stimulus d'une série (1^{ère} unité tonale), et à leur maximum à l'issue du dernier stimulus de la série. Il est à noter que pour ce dernier stimulus, qui ne correspond pas à la fin de la phrase nominale, les réponses pertinentes représentent moins de la moitié des réponses possibles. Ce résultat montre la difficulté qu'ont les auditeurs à identifier l'intonation du stimulus antérieurement à la fin de l'énoncé de la phrase. Cependant, l'évolution des réponses pertinentes sur cette partie de l'énoncé va dans le sens attendu.

Par ailleurs un stimulus affirmatif donne lieu à davantage de réponses pertinentes qu'un stimulus interrogatif. Cet écart s'amenuise au cours du dévoilement de la 2^{ème} unité tonale et la différence cesse d'être significative à l'issue de la présentation du stimulus qui représente 60% de l'unité tonale (4^{ème} essai). L'avantage dont semblent bénéficier pour un temps les stimulus affirmatifs peut être interprété comme un biais dans les réponses, issu de l'application d'une stratégie qui consisterait à donner d'emblée une réponse d'identification « A » de préférence à une réponse « Q » en se référant à la

fréquence des deux types d'énoncés dans la langue, quitte à modifier ce choix à l'issue des stimulus suivants. L'examen des réponses à l'issue du premier stimulus montre en effet qu'en l'absence de toute information prosodique discriminante, dans plus de 70% des cas (70,51 %) les participants donnent une réponse « A ».

Analyse des patrons de réponse

Pour pouvoir confirmer l'interprétation d'une réduction progressive de l'écart entre les réponses pour un stimulus affirmatif et un stimulus interrogatif par la prise en compte d'indices prosodiques discriminants, il est nécessaire de considérer non pas le nombre de réponses pertinentes données par le groupe en fonction du pourcentage de stimulus dévoilé, mais d'étudier l'évolution des réponses de chaque participant au cours de la série de stimulus. Il faut considérer pour cela les seize patrons de réponse à six places, donnés par chacun au cours de l'expérience.

Si, pour une série de stimulus, un auditeur a décidé de répondre d'emblée « A » et de maintenir ce type de réponse jusqu'à preuve trouvée dans le stimulus que ce choix ne convient pas et si le stimulus proposé est affirmatif, alors ses réponses sont toutes pertinentes. Son patron de réponse est : « 111111 ». Si le stimulus proposé est interrogatif et que l'auditeur n'a pas découvert d'indice contraire à son choix initial, le patron de ses réponses est « 000000 ». S'il lui a semblé nécessaire de modifier à une ou à plusieurs reprises son choix au cours des essais, le patron de ses réponses est mixte (composé de zéro et de un en nombre variable). Il en est de même pour un auditeur qui aurait décidé de répondre d'emblée « Q » au premier stimulus de la série.

L'examen des 512 patrons de réponse (16 items x 32 participants) (Figure 5) montre que plus d'une fois sur deux (54,88 % des patrons), les participants ne modifient pas leur choix initial au cours de l'audition de la série de stimulus. Ceci les conduit dans la moitié des cas à fournir des réponses non pertinentes. Comme supposé à l'examen des résultats globaux, le choix initial majoritaire consiste à répondre systématiquement plutôt « A » (40,43 %) que « Q » (14,45 %).

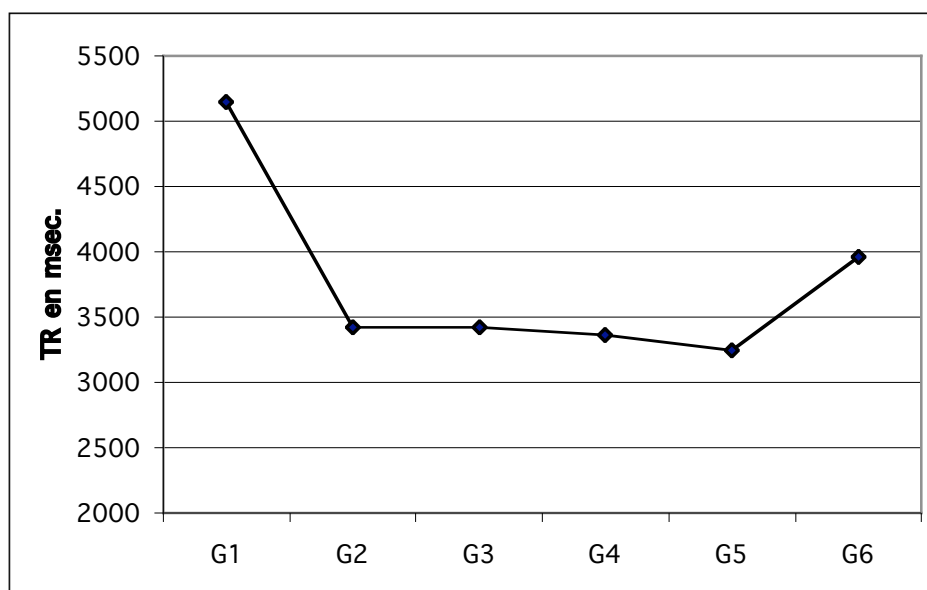


Figure 5 : Distribution des patrons de réponse en fonction de l'intonation du stimulus.

Légende : IA : intonation affirmative ; IQ : intonation interrogative ; RA : réponse systématique « A » ; RQ : réponse systématique « Q » ; RM : réponses mixtes.

Ce résultat souligne la difficulté des auditeurs à déceler les indications prosodiques offertes par les stimulus dans une tâche déconnectée des circonstances ordinaires de la communication. Cependant des 32 participants, personne n'a jamais dévié de son choix initial. Les patrons mixtes, attestant d'une réponse non-systématique aux stimulus proposés (45,12 %), se rencontrent chez tous les participants.

Seul l'examen de ces patrons de réponse peut permettre de fournir des arguments en faveur de l'hypothèse de la capacité de l'auditeur à identifier le pattern de la question antérieurement à l'audition de l'unité tonale porteuse du contour final.

Pour cela, une nouvelle variable dépendante a été dérivée des patrons mixtes, en tenant compte seulement des réponses données au cours du dévoilement de la 2^{ème} unité tonale. Soit des patrons à 5 places correspondant aux réponses aux stimulus G2, G3, G4, G5 et G6.

Ces patrons peuvent être le résultat ou bien d'une fluctuation de la réponse (patron dit « fluctuant » : « 01101 ») ou bien d'une stabilisation de la réponse pertinente au fur et à mesure du dévoilement du stimulus (patron dit « pertinent » : « 01111 »). Dans ce dernier cas, on peut supposer que l'auditeur a prélevé en un point de la série de stimulus (en G3 dans l'exemple), un indice en faveur de l'une des intonations possibles et que la présentation des stimulus suivants conforte son choix.

La figure 6 présente la distribution de ces deux types de patrons en fonction de l'intonation du stimulus. Les patrons fluctuants dominent (66,95 % des

patrons), attestant une fois de plus de la difficulté de la tâche. Ces patrons concernent surtout les stimulus affirmatifs, alors que les patrons pertinents (33,05 %) concernent surtout les stimulus interrogatifs. Ce dernier résultat va dans le sens attendu : bien qu'il soit difficile de le faire, les auditeurs se sont montrés capables de déterminer la modalité de phrase en cours d'audition de l'unité tonale intermédiaire.

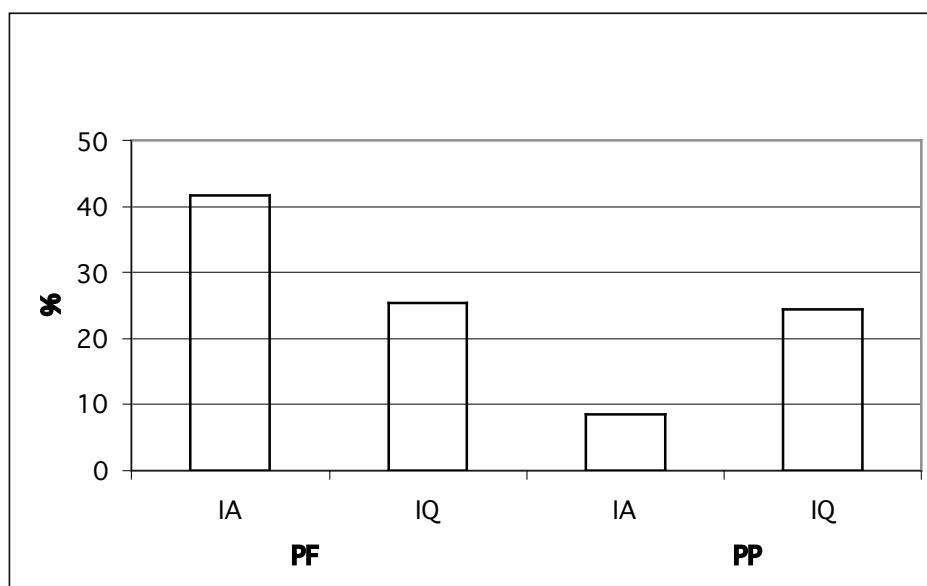


Figure 6 : Distribution des patrons à réponses mixtes en fonction de l'intonation du stimulus.

Légende : IA : intonation affirmative ; IQ : intonation interrogative ; PF : patron mixte fluctuant ; PP : patron mixte pertinent.

Les patrons pertinents donnés en réponse à un stimulus interrogatif (56 patrons) sont environ trois fois plus nombreux que les patrons pertinents donnés en réponse à un stimulus affirmatif (20 patrons). En outre, ces patrons ont été donnés seulement par 14 participants en réponse à un stimulus affirmatif et par 26 participants en réponse à un stimulus interrogatif.

Les 56 patrons donnés en réponse aux stimulus interrogatifs représentent 21,9 % des 256 patrons où ce type de réponse pouvait être attendu ce qui est une performance relativement faible. En revanche, plus de 80 % des participants (81,25 %) se sont montrés capables de le donner au moins une fois. Onze participants n'ont donné ce patron qu'une seule fois au cours de l'expérience, contre quinze qui l'ont donné plusieurs fois (le nombre maximum de patrons de ce type rencontré alors chez un auditeur est : 5).

Point d'isolement

La présence de patrons pertinents donnés en réponse aux stimulus interrogatifs par une large majorité d'auditeurs permet de poursuivre plus

avant l'analyse afin d'estimer quelle est la partie de l'information prosodique entrante qui leur a été nécessaire pour dépasser le seuil d'une réponse au hasard. Pour cette dernière analyse, un seul patron de réponse a été associé à chacun des 26 participants ayant donné au moins une fois un patron pertinent en réponse à un stimulus interrogatif au cours de l'expérience. Ce patron correspond pour certains, au seul patron pertinent donné et pour les autres, à un patron moyen. Ainsi que le montre la figure 7 la fréquence de la réponse pertinente dépasse le hasard à partir du 4^{ème} essai après que 60 % de la 2^{ème} unité tonale a été entendue.

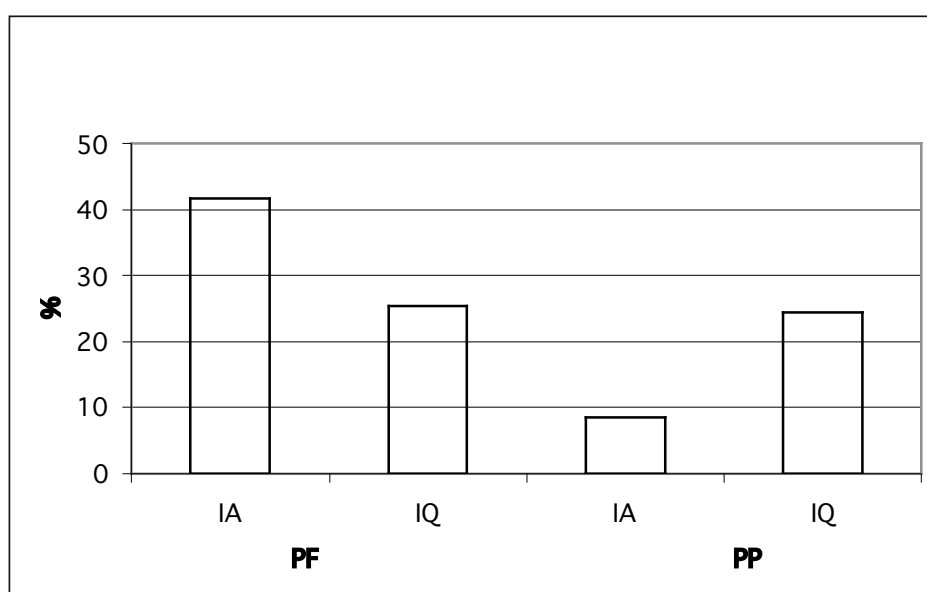


Figure 7 : Patrons pertinents : évolution de la fréquence des réponses pertinentes au cours du dévoilement de la 2^{ème} unité tonale.

Légende : G2 : 20 % de la 2^{ème} unité tonale ; G3 : 40 % de la 2^{ème} unité tonale ; G4 : 60 % de la 2^{ème} unité tonale ; G5 : 80 % de la 2^{ème} unité tonale ; G6 : 100% de la 2^{ème} unité tonale.

Les données recueillies permettent de calculer pour les 26 participants quelle est la certitude moyenne associée à la réponse en ce point. Au 4^{ème} essai celle-ci est encore inférieure au milieu de l'échelle (Figure 8).

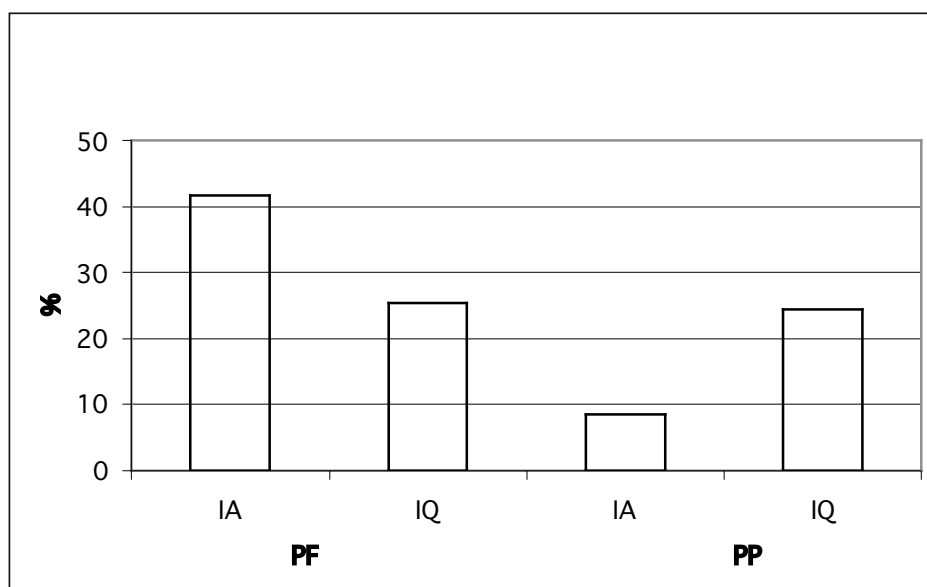


Figure 8 : Patrons pertinents : évolution du degré moyen de la certitude au cours du dévoilement de la 2^{ème} unité tonale. (1 = "pas sûr du tout" et 9 = "tout à fait sûr"; N = 26 participants).

Légende : G2 : 20 % de la 2^{ème} unité tonale ; G3 : 40 % de la 2^{ème} unité tonale ; G4 : 60 % de la 2^{ème} unité tonale ; G5 : 80 % de la 2^{ème} unité tonale ; G6 : 100% de la 2^{ème} unité tonale.

Il est à noter que ces mêmes auditeurs ne sont que moyennement sûrs de la réponse qu'ils donnent à l'issue du dernier essai.

DISCUSSION

L'étude a trouvé son point de départ dans l'analyse du marquage formel de la modalité interrogative par la prosodie. Cette analyse fait état d'abaissements successifs de la hauteur, antérieurs au contour final, observables dès que la composition des unités intonatives dépasse deux unités tonales. Ces abaissements constitueraient une propriété phonologique de la question en français. Partant de là, l'objectif était de mettre expérimentalement en évidence leur valeur distinctive pour l'auditeur, lors de la reconnaissance de la modalité de phrase.

D'un point de vue psychologique, reconnaître un élément de langage parlé suppose la mise en œuvre d'un processus de comparaison entre deux représentations : une représentation d'entrée qui recode de façon interne l'information issue du signal acoustique et une représentation mentale préalable de la forme à reconnaître. Pour que l'élément concerné soit reconnu, la représentation d'entrée élaborée par l'auditeur doit être discriminante en ce qui concerne l'ensemble des appariements possibles.

Le paradigme du dévoilement graduel, en limitant la quantité d'information entrante, permet par le contrôle des propriétés temporelles du signal acoustique ainsi exercé, de déterminer à partir de quand des informations

pertinentes deviennent disponibles pour alimenter le processus de comparaison et conduire à une reconnaissance effective. La mise en œuvre de ce paradigme au service de l'étude du traitement exclusif de l'information prosodique, a fait surgir un certain nombre de difficultés pour les auditeurs. Cet article tenait à les souligner et à les illustrer.

La première a trait à la nature même de l'activité de reconnaissance. Le caractère strictement perceptif de la reconnaissance fait l'objet de désaccords entre les chercheurs. Pour certains, outre les informations dérivées de l'entrée sensorielle, la reconnaissance peut s'alimenter des informations issues des connaissances de l'auditeur (connaissances lexicales, syntaxiques, sémantiques et pragmatiques). Cela semble avoir été le cas lors du sondage préalable à l'expérience (qui dévoilait les phrases, unité tonale par unité tonale). Les phrases nominales choisies comme vecteurs de l'information prosodique sont interprétées dans un contexte communicatif donné comme l'énoncé de phrases complètes (assertives ou interrogatives). Ces phrases se sont avérées cesser parfois d'être considérées comme des phrases achevées quand, désolidarisées de tout contexte discursif, elles ont été présentées isolément. Il semble que hors de tout contexte communicatif, malgré les indications explicites données dans la consigne, la représentation de la phrase convoquée par certains auditeurs pour la comparaison ait été celle de la phrase verbale. Malgré cela, les résultats du sondage montraient que majoritairement les auditeurs considéraient les énoncés proposés comme des phrases et que la modalité de phrase, bien que peu fréquemment, pouvait être reconnue antérieurement à l'audition de la fin de l'énoncé (à l'issue de la dernière unité tonale non-finale).

La seconde difficulté rencontrée a trait au processus de comparaison lui-même. Il s'agit de comparer la représentation de segments de phrases entendus, brefs et transitoires, à la représentation mentale de patrons prosodiques.

Une raison de la difficulté des participants, rencontrée lors du sondage, peut avoir sa source dans le mode de réponse choisi. Contrairement à la pratique en usage dans la tâche de dévoilement graduel, un rappel visuel de l'alternative n'était pas présenté lors de l'invitation à la réponse. Or on ne peut pas exclure que la représentation préalable à laquelle doit être comparé le segment entendu trouve un appui solide dans la sub-vocalisation issue de la lecture de la version écrite et ponctuée des phrases alternatives. Ceci a conduit à modifier en conséquence le mode de réponse de la présente expérience.

Une autre raison de la difficulté des auditeurs, rencontrée lors de l'expérience, peut avoir sa source dans une caractéristique inhérente à la tâche de dévoilement graduel. Il lui est parfois reproché d'induire chez

l'auditeur des stratégies s'apparentant à la devinette. C'est le cas notamment quand le stimulus présenté ne contient pas encore d'information discriminante. Dans l'expérience, les auditeurs ont fréquemment eu recours à une réponse par défaut qui était de considérer a priori que le stimulus provenait plutôt d'une phrase affirmative que d'une phrase interrogative. En outre la tâche était rendue particulièrement ardue du fait que les auditeurs, hormis lors de la délivrance de la consigne, n'entendaient jamais les phrases entièrement dévoilées. Ils ne pouvaient donc pas recevoir en cours d'expérience d'indications prosodiques issues du dévoilement complet des phrases qui auraient en retour conforté ou bien contredit leur choix. Ceci peut expliquer qu'ils aient souvent persévéré dans leur réponse initiale jusqu'au dernier stimulus présenté.

Tous les auditeurs ont montré des persévérations dans l'emploi de la réponse stratégique initiale, mais tous également ont montré qu'ils étaient attentifs aux stimulus. En premier lieu, la mesure des temps mis pour donner une réponse d'identification, ce qui ne se pratique pas d'ordinaire dans le paradigme de dévoilement graduel, montre que les auditeurs sont moins rapides à donner leur réponse pour les deux stimulus extrêmes de la série. Ces stimulus dévoilent complètement une ou deux unités tonales ainsi qu'une séquence phonique porteuse de sens (Art + N ou Art + N + prép. + N). Dans ces stimulus, les informations prosodiques et les informations segmentales sont complètement assemblées. Ces stimulus représentent peut-être des points privilégiés du décours temporel du signal pour effectuer la comparaison. En second lieu, lorsque la réponse en cours d'item n'a pas été maintenue systématiquement identique au choix initial, l'évolution des réponses demeure le plus souvent fluctuante. Mais, bien que le phénomène se révèle minoritaire et que les auditeurs ne soient que moyennement sûrs des réponses qu'ils donnent, une partie des réponses s'organise selon un patron compatible avec la reconnaissance de la prosodie du signal d'entrée. Un des participants ayant fourni de tels patrons, interrogé à l'issue de la session expérimentale, explique : « quand c'est une question, c'est comme si ça allait monter. ».

Il reste à conduire une nouvelle expérience qui offre à la fois des conditions optimales pour la comparaison et qui contrôle les caractéristiques du signal acoustique dévoilé à chaque pas. Sur le premier point, l'option choisie sera de conserver le mode de réponse mis au point pour la présente expérience et de poursuivre le dévoilement jusqu'à audition complète des phrases. Sur le second point, l'option sera de constituer les stimulus, non pas à partir de l'enregistrement brut des phrases, mais sur la base de phrases où l'évolution de la fréquence fondamentale aura été vérifiée et calibrée point par point.

Bibliographie

COHEN, J., MAC WHINNEY, B., FLATT, M., & PROVOST, J. (1993). An interactive graphical system for designing and controlling experiments in the psychological laboratory using MacIntosh computers. *Behavior Methods, Research, Instruments and Computer*, 25, 257-271.

DI CRISTO, A. & HIRST, D. J. (1993). Prosodic regularities in the surface structure of French questions. *Proceedings of an European Speech Communication Association Workshop on Prosody*. Lund: Lund University, Department of Linguistics.

DI CRISTO, A. (1998). Intonation in French. In D., Hirst et A. Di Cristo (Eds.) *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press, 195-218.

DI CRISTO, A. (2000). Interpréter la prosodie. XXIIIèmes Journées d'Etude sur la Parole, juin 2000, Aussois, France, p. 13-29.

GROSJEAN, F. & HIRT, C. (1996). Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. *Language and Cognitive Processes*, 11, (1/2), 107-134.

GROSJEAN, F. (1980a). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, 28, 267-283.

GROSJEAN, F. (1980b). How long is the sentence? Prediction and prosody in the on-line processing of language, *Linguistics*, 21, 501-529.

GROSJEAN, F. (1996). Gating. *Language and Cognitive Processes*, 11, 597-604.

HIRST, D. ET DI CRISTO, A. (1998). A survey of intonation systems. In D., Hirst et A. Di Cristo (Eds.) *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press, 1-44.

HIRST, D.J. & DI CRISTO, A. (1984). French intonation: A parametric approach. *Die Neueren Sprachen*, 83, 554-569.

HIRST, D.J. (1988). Tonal units as phonological constituents: The evidence from French and English intonation. In H. Van der Hulst & N. Smith (Eds.) *Autosegmental Studies in Pitch Accent*. Dordrecht : Foris, 151-165.

LE GOFFIC, P. (1993). *Grammaire de la phrase française*. Paris : Hachette.

SOULIER, I. (1995). *Rôle de l'intonation dans la perception des énoncés nominaux interrogatifs: peut-on identifier une question avant la fin d'un énoncé?* Note de recherche non publiée. Aix-en-Provence : Université de Provence.

ANNEXE 1

liste des phrases

Le syntagme mis en italiques, qui est le support du contour final (dernière unité tonale) ne figure pas dans les stimulus.

Phrases pour les stimulus test :

les ragots du bureau *d'Amélie*
 la maison du voisin *d'Eugénie*
 l'agonie du géant *d'Arabie*
 les relents des égouts *de Dehli*
 la durée du visa *d'Amédée*
 le ronron du minet *du logis*
 le ballon du gamin *d'Hélène*
 le début du roman *de Zola*
 le réseau des amis *du Jura*
 le guidon du vélo *d'Amanda*
 les ébats des amants *d'Ibiza*
 le venin du naja *d'Ouganda*
 le radeau du marin *d'Oléron*
 les billets du magot *d'Odilon*
 les manies du valet *du baron*
 les navets du ragoût *d'Honorin*

Phrases pour les stimulus d'entraînement :

les rabats du gilet *d'Ugolin*
 l'enjeu des débats *des ruraux*
 la magie du génie *d'Aladin*

ANNEXE 2

Consigne

Vous allez entendre des phrases tronquées (c'est-à-dire des phrases dont on a enlevé quelques mots de la fin). Les phrases d'origine avaient des intonations variées. C'étaient des AFFIRMATIONS ou des QUESTIONS, comme on en rencontre dans la conversation courante.

Par exemple, vous rencontrez un ancien compagnon de lycée qui vous demande :

- qu'est-ce que tu fais cette année ?

- je suis étudiant à la fac des Lettres d'Aix.(A)

- étudiant à la fac des Lettre d'Aix ? (Q) Et ça te plaît ?

Ecoutez, on va l'entendre...

[écoute du dialogue]

Dans l'expérience, on a pris des phrases comme : " étudiant à la fac des Lettres d'Aix ", c'est-à-dire des phrases sans verbe, et on a enlevé le dernier groupe de mots. Par exemple ici, on a enlevé " d'Aix ". Ce qui donne pour cet exemple les deux stimulus suivants :

[écoute des stimulus]

Votre tâche sera de deviner si les phrases entières dont les stimulus sont issus sont des AFFIRMATIONS ou des QUESTIONS. Comme on aurait pu vous demander de le faire en vous présentant l'un des deux "étudiant à la fac des lettres" et en vous demandant de dire si ce stimulus provient de la phrase "A" ou de la phrase "Q".

ATTENTION ! Les stimulus que vous allez entendre vont être dévoilés graduellement, c'est-à-dire que vous allez d'abord entendre : un premier stimulus qui comporte les mots du début de la phrase, par exemple : " étudiant à la fac ". Puis un deuxième stimulus qui comporte les mots déjà entendus, plus un petit bout des mots qui suivent. Et ainsi de suite jusqu'à entendre complètement le dernier mot du stimulus.

A chaque fois que vous entendrez un stimulus, vous devrez dire de quelle phrase A ou Q ce stimulus provient. Pour cela, vous appuierez sur les touches étiquetées sur le clavier : " A " et " Q ".

Comme vous ne pourrez pas toujours être certain de votre réponse, vous exprimerez également le degré de certitude que vous accordez à cette réponse. Vous exprimerez votre certitude au moyen des chiffres " 1 " à " 9 " du pavé numérique où : " 1 " signifiera " pas du tout sûr ", " 9 " signifiera " tout à fait sûr ".

Notez aussi : à la fin de l'expérience il vous sera demandé quels sont les INDICES qui vous ont permis de décider qu'il s'agissait d'un stimulus issu d'une AFFIRMATION ou bien d'une QUESTION.

On va d'abord faire quelques essais pour que vous soyez familiarisé avec le dispositif de l'expérience, puis on passera à l'expérience proprement dite.

Vous avez tout votre temps pour donner vos réponses.

Pour commencer l'entraînement, appuyez sur la barre espace